DØ Note 4878-CONF v2.1

Search for B_s^0 oscillations at DØ using $B_s^0 \to D_s^- \mu^+ X$ ($D_s^- \to K^{*0} K^-$) decays

The DØ Collaboration
URL http://www-d0.fnal.gov
(Dated: July 16, 2005)

The $B_s^0 \to D_s^- \mu^+ X$ decay into the $D_s^- \to K^{*0} K^-$ ($K^{*0} \to K^+ \pi^-$) final state was reconstructed at DØ using ~ 610 pb^{-1} of data. A search for B_s^0 oscillations was performed using opposite side flavor tagging algorithms. A 95% confidence level limit on the oscillation frequency $\Delta m_s > 4.9$ ps^{-1} and a sensitivity of 7.4 ps^{-1} were obtained. A combination with the $B_s^0 \to D_s^- \mu^+ X$ ($D_s^- \to \phi \pi^-$) decay mode improves the result for the limit to $\Delta m_s > 7.3$ ps^{-1} and sensitivity to 9.5 ps^{-1}.

Preliminary Results for Summer 2005 Conferences

I. INTRODUCTION

Mixing is the process whereby some neutral mesons change from their particle to their anti-particle state, and vice versa. This kind of oscillation of flavor eigenstates into one another was first observed in the K^0 meson system. It has since then been seen for B mesons, first in a mixture of B_d^0 and B_s^0 by UA1 and then in B_d^0 mesons by ARGUS [1]. The frequency of the oscillation is proportional to the small difference in mass between the two eigenstates, Δm, and for the $B_d^0 - \bar{B_d^0}$ system, can be translated into a measurement of the CKM element $|V_{td}|$. Δm_d has been precisely measured (the world average is $\Delta m_d = 0.502 \pm 0.007$ ps^{-1}) [2] but large theoretical uncertainties dominate the extraction of $|V_{td}|$ from Δm_d. This problem can be reduced if the B_s^0 mass difference, Δm_s, is also measured. $|V_{td}|$ can then be extracted with better precision from the ratio:

$$\frac{\Delta m_s}{\Delta m_d} = \frac{m(B_s^0)}{m(B_d^0)} \xi^2 \mid \frac{V_{ts}}{V_{td}} \mid^2 \tag{1}$$

where ξ is estimated from Lattice QCD calculations to be $1.15 \pm 0.05^{+0.12}_{-0.00}$. The above has motivated many experiments to search for B_s^0 oscillations, and though a statistically significant signal hasn't been observed yet, a lower limit of $\Delta m_s > 14.4$ ps^{-1} at 95% C.L. [2] has been set. Since this current limit indicates that the B_s^0 oscillations are at least 30 times faster than the B_d^0 oscillations, a B_s^0 mixing measurement is experimentally very challenging. If the Standard Model is correct, then Δm_s is expected from global fits to the unitarity triangle to be in the range $(16.2 - 24.5)$ ps^{-1} at the one standard deviation confidence level [3].

II. DETECTOR DESCRIPTION

The following main elements of the DØ detector are essential for this analysis:

- A magnetic central-tracking system, which consists of a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located within a 2 T superconducting solenoidal magnet;
- A muon system located beyond the calorimetry.

The SMT has $\approx 800,000$ individual strips, with typical pitch of $50 - 80$ μm, and a design optimized for tracking and vertexing capability at $|\eta| < 3$, where $\eta = -\ln(\tan(\theta/2))$. The system has a six-barrel longitudinal structure, each with a set of four layers arranged axially around the beam pipe, and 16 radial disks. The CFT has eight thin coaxial barrels, each supporting two doublets of overlapping scintillating fibers of 0.835 mm diameter, one doublet being parallel to the collision axis, and the other alternating by $\pm 3^\circ$ relative to the axis. Light signals are transferred via clear light fibers to solid-state photon counters (VLPC) that have $\approx 80\%$ quantum efficiency.

The muon system consists of a layer of tracking detectors and scintillation trigger counters before 1.8 T toroids, followed by two additional layers after the toroids. Tracking at $|\eta| < 1$ relies on 10 cm wide drift tubes, while 1 cm mini-drift tubes are used at $1 < |\eta| < 2$.

III. EXPERIMENTAL CONSIDERATIONS

Every mixing measurement is, at its heart, an asymmetry measurement. The asymmetry, $A^{\rm meas}$, as a function of time is written as:

$$A^{\rm meas}(t_{B_s^0}) = \frac{N^{\rm non-osc}(t_{B_s^0}) - N^{\rm osc}(t_{B_s^0})}{N^{\rm non-osc}(t_{B_s^0}) + N^{\rm osc}(t_{B_s^0})} \propto \cos(\Delta m_s t_{B_s^0}) \tag{2}$$

where $N^{\rm non-osc}$ and $N^{\rm osc}$ are the number of non-oscillated (unmixed) and oscillated (mixed) mesons, respectively. This analysis, therefore, is based on the simple observation that we can extract the mass difference, Δm_s, from the period of the oscillation.

The different elements essential to a mixing analysis include (a) reconstruction of final states, (b) proper time determination, (c) determination of whether a B_s^0 meson has mixed or not and the success/error rate associated with this estimation, and (d) fitting for Δm_s.

Two semileptonic $B_s^0 \to D_s^- \mu^+ X$ [12] decays are currently being studied at DØ the $D_s^- \to \phi\pi^-$ mode [4] and the $D_s^- \to K^{*0} K^-$ mode. Details of the reconstruction of the semileptonic $B_s^0 \to D_s^- \mu^+ X$ decay into the $D_s^- \to$

$K^{*0}K^-$ final state using data from the Run II DØ detector are presented in this note. This analysis uses ~ 610 pb^{-1} of data collected during the April 2002 - May 2005 period.

The proper lifetime of the B_s^0 meson, $ct_{B_s^0}$, for semileptonic decays can be written as:

$$ct_{B_s^0} = x^M \cdot K, \text{ where } K \equiv p_T^{D_s^- \mu}/p_T(B_s^0) \text{ and } x^M \equiv (\mathbf{L_{xy}} \cdot \mathbf{p_{xy}^{D_s^- \mu}})/(p_T^{D_s^- \mu})^2 \cdot M_{B_s^0}. \quad (3)$$

K is a correction factor used in semileptonic decays to account for the missing neutrino (and other neutral or non-reconstructed charged particles) and is obtained using Monte Carlo simulations. x^M is the *visible proper decay length* or VPDL, L_{xy} is the axial decay length of the B_s^0 meson, $p_T^{D_s^- \mu}$ is the transverse momentum of the $D_s^- + \mu$ system and $M_{B_s^0}$ is the mass of the B_s^0 meson.

The flavor of the B_s^0 meson at decay is obtained using the charge of the final state lepton (muon); a negative muon corresponds to a b quark, and vice versa. The methods for tagging the initial state can be grouped into two categories: the ones that tag the initial charge of the b quark in the B_s^0 candidate itself (same-side tag), and those that tag the flavor of the other b quark in the event (opposite-side tag). Different techniques have been developed and tested with measurements of Δm_d, and the success rate of the tagging algorithms given by "dilution", $\mathcal{D}$, or purity, η_s, obtained[5]:

$$\mathcal{D} \equiv \frac{\mathrm{N_{correct}} - \mathrm{N_{wrong}}}{\mathrm{N_{correct}} + \mathrm{N_{wrong}}} \; ; \; \eta_s \equiv N_{\mathrm{correct}}/N_{\text{total tagged events}} \text{ and } \mathcal{D} \equiv 2\eta_s - 1 \quad (4)$$

where N_{correct} is the number of correct tags and N_{wrong} is the number of wrong tags. For this analysis we concentrate on the opposite side tagging algorithms since in that case the purities obtained from the Δm_d measurements can be used directly (the charge of the b quark on the opposite side should not be affected by whether there is a B_d^0 or a B_s^0 meson on the reconstructed side).

The fitting procedure for Δm_s involves calculating an expected asymmetry, A^e, keeping in mind the K factor correction, the VPDL resolution and the fact that there can be contamination of the B_s^0 sample by mesons that either do not mix ($B^\pm$) or mix at a different rate (B_d^0). Then a time dependent asymmetry between unmixed and mixed mesons is obtained. This is done by producing D_s^- mass distributions for different VPDL bins, for both the unmixed and mixed event samples and then determining the numbers of unmixed and mixed B mesons for each bin by fitting the distributions to functions describing the signal and background contributions.

The experimental observable, asymmetry A_i^{meas}, in each VPDL bin, i, is defined as:

$$A_i^{\mathrm{meas}} = \frac{N_i^{non-osc} - N_i^{osc}}{N_i^{non-osc} + N_i^{osc}} \quad (5)$$

where $N_i^{non-osc}$ is the number of events tagged as "non-oscillated" and N_i^{osc} is the number of events tagged as "oscillated". A technique called the *amplitude fit method* [6] can then be used to study the B_s^0 oscillations. According to this method, the frequency of the oscillation is not taken to be a free parameter but is instead fixed to a "test" value ω. An auxiliary parameter, the amplitude $\mathcal{A}$ of the oscillating term, is introduced, and left free in the fit. The fitted values of $\mathcal{A}$ as a function of ω can then be determined from the minimization of a $\chi^2(\mathcal{A})$ defined as:

$$\chi^2(\mathcal{A}) = \sum_i \frac{(A_i^{\mathrm{meas}} - A_i^e(\mathcal{A}))^2}{\sigma^2(A_i^{\mathrm{meas}})}. \quad (6)$$

When the test frequency is much smaller than the true frequency ($\omega \ll \Delta m_s$), the expected value for the amplitude is $\mathcal{A} = 0$, while at the true frequency ($\omega = \Delta m_s$) the expectation is $\mathcal{A} = 1$. All values of the test frequency ω for which $\mathcal{A} + 1.645\sigma_{\mathcal{A}} < 1$ can be excluded at 95% C.L..

IV. RECONSTRUCTION AND EVENT SELECTION

The selection of the channel $D_s^- \to K^{*0}K^-$ from the $B_s^0 \to D_s^- \mu^+ X$ decay is described here. The muon in the event was identified using the standard DØ algorithm [7]. It was required to have $p_T > 2$ GeV/c and $p > 3$ GeV/c, to have hits both in the CFT and SMT and to have at least two measurements in the muon chambers. All tracks in an event were clustered into jets using the DURHAM clustering algorithm with the cutoff parameter of 15 GeV/c [8]. Three additional charged tracks were required to be from the same jet as the muon, to have hits both in the SMT and CFT and to have total charge equal to one in magnitude and opposite to the charge of the muon. The particles were assigned the masses of kaons (K_1 and K_2) and pion requiring the charge combination $\mu^+ K_1^+ K_2^- \pi^-$ or its charge

conjugate. The transverse momenta were required to be: $p_T(K_1) > 0.9$ GeV/c, $p_T(K_2) > 1.8$ GeV/c and $p_T(\pi) > 0.5$ GeV/c, assuming that K_1 comes from the $K^{*0} \to K^+\pi^-$ decay.

For each particle, the transverse[13] ϵ_T and longitudinal[14] ϵ_L projections of track impact parameter with respect to the primary vertex together with the corresponding errors ($\sigma(\epsilon_T)$, $\sigma(\epsilon_L)$) were computed. The combined significance $(\epsilon_T/\sigma(\epsilon_T))^2 + (\epsilon_L/\sigma(\epsilon_L))^2$ was required to be greater than 4 for K_1 and K_2, while there was no cut on the significance of the pion.

Three charged particles were required to come from the same D_s^- vertex with the χ^2 of the vertex fit satisfying $\chi^2 < 16$. The D_s^- candidate produced by their combination was required to have a common B vertex with the muon with the χ^2 of the vertex fit such that $\chi^2 < 9$. The mass of the $\mu^+ D_s^-$ system was required to be in the range $2.6 < M(\mu^+ D_s^-) < 5.4$ GeV/c^2. The distance d_T^D in the axial plane between the D_s^- vertex and the primary interaction point was required to satisfy $d_T^D/\sigma(d_T^D) > 4$. The angle α_T^D between the momentum direction of the D_s^- candidate and the direction from the primary to the D_s^- vertex in the axial plane was required to fulfill the condition $\cos(\alpha_T^D) > 0.9$.

If the distance d_T^B between the primary and B_s^0 vertex in the axial plane exceeded $4 \cdot \sigma(d_T^B)$, the angle α_T^B between the B_s^0 momentum and the direction from the primary to B_s^0 vertex in the axial plane was required to satisfy the condition: $\cos(\alpha_T^B) > 0.95$. The distance d_T^B was allowed to be greater than d_T^D, provided that the distance between the B_s^0 and D_s^- vertices, d_T^{BD}, was less than $2 \cdot \sigma(d_T^{BD})$. Additionally, the K^{*0} mass was required to be in the range $0.82 < M(K_1\pi) < 0.95$. The final event samples were then selected using the likelihood ratio method, described below.

It is assumed that a set of discriminating variables $x_1, ...x_n$ can be constructed for a given event. It is also assumed that the probability density functions $f^s(x_i)$ for the signal and $f^b(x_i)$ for the background can be built for each variable x_i. The combined tagging variable y is defined as:

$$y = \prod_{i=1}^{n} y_i; \quad y_i = \frac{f_i^b(x_i)}{f_i^s(x_i)}. \tag{7}$$

A given variable x_i can be undefined for some events. In this case, the corresponding variable y_i is set to one. The selection of the signal is obtained by applying the cut on $y < y_0$. For uncorrelated variables $x_1, ...x_n$, the selection using the combined variable y gives the best possible tagging performance, i.e., maximal signal efficiency for a given background efficiency.

The following discriminating variables were used:

- Helicity angle, defined as the angle between the D_s^- and K_1 momenta in the $(K_1\pi)$ center of mass system;
- Isolation, computed as $\mathrm{Iso} = \mathrm{p^{tot}}(\mu\mathrm{D_s})/(\mathrm{p^{tot}}(\mu\mathrm{D_s}) + \sum \mathrm{p_i^{tot}})$. The sum $\sum p_i^{tot}$ was taken over all charged particles in the cone $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.5$, where $\Delta\eta$ and $\Delta\phi$ are the pseudorapidity and the azimuthal angle with respect to the (μD_s) direction. The μ^+, K_1, K_2 and π^- were not included in the sum;
- $p_T(K_2)$;
- Invariant mass, $M(\mu^+ D_s^-)$;
- χ^2 of the D_s^- vertex fit;
- $M(K_1\pi)$.

The probability density functions were constructed using real data events. For each channel, three bands B_1, B_2 and S were defined as:

$$\begin{aligned} B_1 &: 1.75 < M(D_s^-) < 1.79 \text{ GeV}/c^2, \\ B_2 &: 2.13 < M(D_s^-) < 2.17 \text{ GeV}/c^2, \\ S &: 1.92 < M(D_s^-) < 2.00 \text{ GeV}/c^2. \end{aligned}$$

The background probability density function for each variable was constructed using events from the B_1 and B_2 bands. The signal probability density function was constructed by subtracting the background, obtained as a sum of distributions in the B_1 and B_2 bands, from the distribution of events in band S. The final selection of events for the analysis was done by applying a cut on the combined variable $\log_{10} y < 0.16$. This cut was selected by requiring the maximal value of $S/\sqrt{S + B_1 + B_2}$. Figure 1 shows the $-\log_{10} y$ distribution for both signal and background.

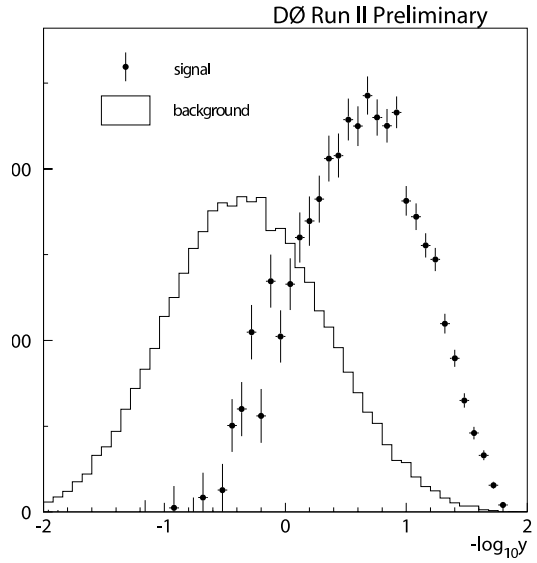

FIG. 1: $-\log_{10} y$ distribution for both signal and background.

A. Mass fitting procedure

Figure 2 shows the D_s^- invariant mass distribution after all the selection cuts. A Gaussian was used to describe the $D_s^- \to K^{*0}K^-$ signal and an exponential was used to model the combinatorial background. The mean and width (σ) of the signal Gaussian were obtained from the final fit. Additionally, to extract an accurate normalization, the following potential reflections were studied and parameterized in the total fit:

- $D^- \to K^+\pi^-\pi^-$ ($\mathcal{B} = 9.2 \pm 0.6\%$):
  A Monte Carlo sample for the Dalitz decay which took into account all the underlying amplitudes - $K^{*0}(892)\pi^-$, $K^{*0}(1430)\pi^-$and $K^{*0}(1680)\pi^-$ resonances and $K^+\pi^-\pi^-_{\text{non resonant}}$ - and the interference between them was generated and used for the study. The mode which poses the biggest background is $K^{*0}(892)\pi^-$ ($\mathcal{B} = 1.30 \pm 0.13\%$), where the pion could be misidentified as a kaon. The other resonances should have a smaller impact since we imposed a cut on the mass of the $K^+\pi^-$ combination. Given that we accept all $K^+\pi^-$ combinations in the K^{*0} mass region, the non-resonant contribution could also make an impact since its branching fraction is quite large ($\mathcal{B} = 8.8 \pm 0.9\%$). However, the mass region occupied by the K^{*0} is much smaller than the phase space available to the $K^+\pi^-$ combinations, and thus the impact should be small. Moreover, the helicity cut should further reduce the non-resonant contribution. Two Gaussians were used to model the decay, and the means and widths of these Gaussians were extracted from this fit and then used to fit the data. When fitting data the means of these two Gaussians were fixed relative to the D_s^- mean and the two widths were allowed to be scaled by a single factor.

- $D^- \to K^{*0}K^-$ ($\mathcal{B} = 2.9 \pm 0.4 \times 10^{-3}$):
  This decay has a final state identical to the $D_s^- \to K^{*0}K^-$ signal. Its contribution was parameterized by a Gaussian with its mean and sigma fixed relative to the D_s^- Gaussian. The number of $D^- \to K^{*0}K^-$ events was obtained from the final fit. The reconstruction and candidate selection efficiencies from Monte Carlo along with the branching fractions were used to estimate the ratio of $D^- \to K^+\pi^-\pi^-$ and $D^- \to K^{*0}K^-$ events and this ratio was used as a fixed parameter when fitting the D_s^- mass distribution in data.

The final fit in data gave an estimated 18780 ± 782 $D_s^- \to K^{*0}K^-$ signal events centered at 1.964 GeV/c^2, and a width of 28 MeV. A total of 3233 ± 208 $D^- \to K^{*0}K^-$ events and 14112 ± 910 $D^- \to K^+\pi^-\pi^-$ events were obtained.

V. INITIAL-STATE TAGGING

The second B meson (or baryon) in the event was used to tag the initial flavor of the reconstructed B_s^0 meson. The tagging technique utilized information from identified leptons (muons and electrons) and reconstructed secondary

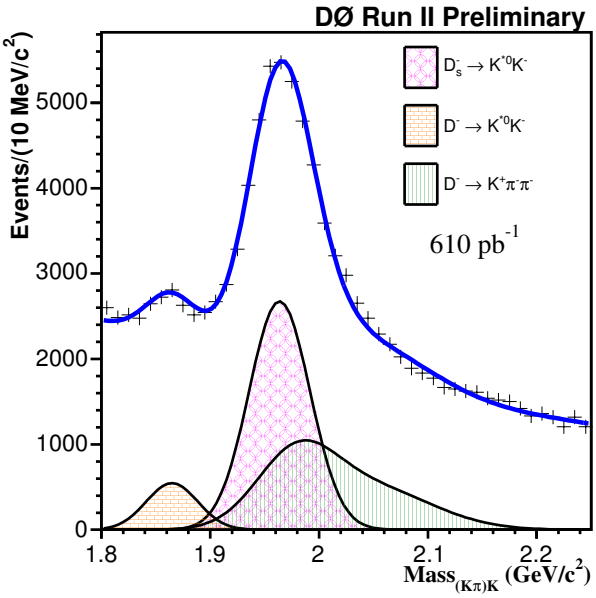

FIG. 2: Distribution of the mass of $D_s^- \to K^{*0}K^-$ candidates. There is a clear $D_s^- \to K^{*0}K^-$ signal peak at 1.964 GeV/c^2. The various reflections and their individual contribution to the total fit are given by the legend description.

vertices. For reconstructed $B_s^0 \to D_s^- \mu^+ X$ decays both leptons having the same sign would indicate that one B hadron had oscillated while opposite signs would indicate that neither (or both) had oscillated. The tagging information from the leptons and secondary vertices was combined in an optimal way and the details of the combination can be found in Ref. [5]. The B_d^0 mixing frequency, Δm_d, was measured to be $\Delta m_d = 0.501 \pm 0.030(stat.)$ ps^{-1} and was found to be in good agreement with the world average of $\Delta m_d = 0.502 \pm 0.007$ ps^{-1} [2]. The dilution for the B_d^0 and $B^\pm$ mesons was determined to be 0.384 ± 0.013 and the corresponding purity value, $\eta_s = 0.692$, was used as an input to the B_s^0 mixing fit.

Figure 3 shows the invariant mass distribution of tagged D_s^- candidates. The fit returned a total of 2247 ± 316 D_s^- tagged candidates indicating a 12% tagging rate or efficiency.

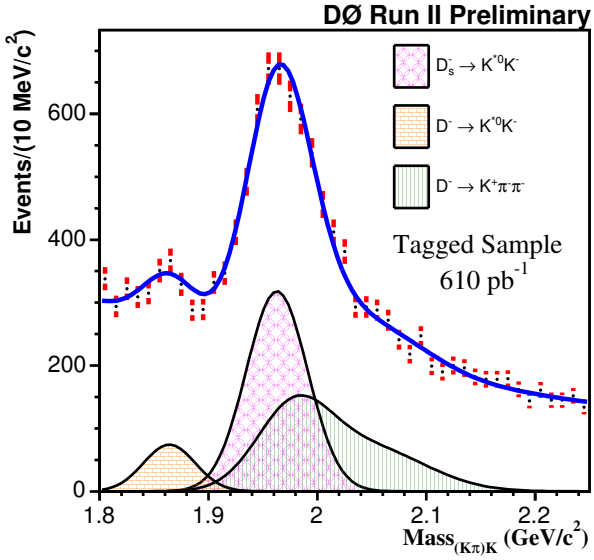

FIG. 3: Distribution of the mass of tagged $D_s^- \to K^{*0}K^-$ candidates.

VI. EXPECTED ASYMMETRY

The D_s^- sample is composed mostly of B_s^0 mesons with some contributions from B_u and B_d^0 mesons. Different species of B mesons behave differently with respect to oscillations; neutral B_d^0 and B_s^0 mesons oscillate while charged B_u mesons do not oscillate. Moreover, contributions of b-baryons to the sample composition are expected to be small and so are neglected. For a given type of B-hadron (i.e. d, u, s), the distribution of the visible proper decay length x for non-oscillated ("non-osc") and oscillated ("osc") events is given by:

$$n_s^{non-osc/osc}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot [1 \pm (2\eta_s - 1)\cos(\Delta m_s \cdot Kx/c)], \quad (8)$$

$$n_{DsDs}^{non-osc}(x) = n_{DsDs}^{osc}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5, \quad (9)$$

$$n_u^{non-osc}(x) = \frac{K}{c\tau_{B_u}} \exp\left(-\frac{Kx}{c\tau_{B_u}}\right) \cdot (1 - \eta_s), \quad (10)$$

$$n_u^{osc}(x) = \frac{K}{c\tau_{B_u}} \exp\left(-\frac{Kx}{c\tau_{B_u}}\right) \cdot \eta_s,$$

$$n_d^{non-osc/osc}(x) = \frac{K}{c\tau_{B_d}} \exp\left(-\frac{Kx}{c\tau_{B_d}}\right) \cdot 0.5 \cdot [1 \mp (2\eta_s - 1)\cos(\Delta m_d \cdot Kx/c)] \quad (11)$$

where K is the K-factor as described before, and τ is the lifetime of the B-hadron taken from the PDG [2]. The $D_s^\pm$ charge has different correlations with the b-quark flavor in the B_u or B_d^0 decays with respect to the B_s^0 semileptonic decays, and Eqs. 10 and 11 take this into account.

The transition to the measured VPDL, x^M, is achieved by the integration over the K-factors and resolution functions:

$$.N_{(d,u,s),\ j}^{osc,\ non-osc}(x^M) = \int dx\ Res_j(x - x^M, x) \cdot Eff_j(x) \cdot \theta(x) \int dK\ D_j(K) \cdot n_{(d,u,s),\ j}^{osc,\ non-osc}(x, K). \quad (12)$$

$Res_j(x - x^M, x)$ is the detector resolution of the VPDL and $Eff_j(x)$ is the reconstruction efficiency for a given decay channel j of this type of B meson. Both are determined from Monte Carlo simulations.

The expected number of oscillated/non-oscillated events in the i-th bin of VPDL is equal to

$$N_i^{e,osc/non-osc} = \int_i dx^M \left(\sum_{f=u,d,s} \sum_j (Br_j \cdot N_{f,\ j}^{osc/non-osc}(x^M)) \right). \quad (13)$$

The integration $\int_i dx^M$ is taken over a given interval i, the sum $\sum_j$ is taken over all decay channels $B \to \mu^+ \nu \mathrm{D}^{*-} X$ and Br_j is the branching ratio of a given channel j.

Finally, the expected value of the asymmetry, A_i^e, for the i-th VPDL bin is given by:

$$A_i^e(\Delta m, \eta_s) = \frac{N_i^{e,non-osc} - N_i^{e,osc}}{N_i^{e,non-osc} + N_i^{e,osc}}. \quad (14)$$

A. Inputs to A_i^e

We have used the following measured parameters for B mesons from the PDG [2] as inputs to the fitting procedure: $c\tau_{B^+} = 501\ \mu$m, $c\tau_{B^0} = 460\ \mu$m, $c\tau_{B_s} = 438\ \mu$m, and $\Delta \mathrm{m_d} = 0.502\ \mathrm{ps}^{-1}$. The latest PDG values were also used to determine the branching fractions of decays contributing to the $\mathrm{D_s^-}$ sample. For those branching fractions not given in the PDG, we used the values provided by the event generator EvtGen [9] since this code was developed specifically for the simulation of B decays and is motivated by theoretical considerations.

Taking into account the corresponding branching rates and reconstruction efficiencies, we determined the following contributions to our signal region from the different processes:

- $\mathrm{B_s^0} \to \mu^+ \nu \mathrm{D_s^-} : 22.8 \pm 1.7\%$;

- $\mathrm{B_s^0} \to \mu^+\nu \mathrm{D_s^{-*}} \to \mu^+\nu \mathrm{D_s^-}$: $55.1 \pm 4.0\%$;
- $\mathrm{B_s^0} \to \mu^+\nu \mathrm{D_{s0}^{*-}} \to \mu^+\nu \mathrm{D_s^-}$: $1.2 \pm 0.1\%$;
- $\mathrm{B_s^0} \to \mu^+\nu \mathrm{D_{s1}^{'-}} \to \mu^+\nu \mathrm{D_s^-}$: $3.0 \pm 0.2\%$;
- $\mathrm{B_s^0} \to \tau^+\nu \mathrm{D_s^-}$; $\tau \to \mu$: $1.6 \pm 0.5\%$;
- $\mathrm{B_s^0} \to \mathrm{D_s^+ D_s^-} X$; $\mathrm{D_s^-} \to \mu\nu X$: $(4.2^{+1.4}_{-3.3})\%$;
- $\mathrm{B_s^0} \to D_s D X$; $D \to \mu\nu X$: $0.9 \pm 0.3\%$;
- $B^+ \to \mathrm{D D_s^-} X$; $\mathrm{D} \to \mu\nu X$: $5.6 \pm 1.9\%$;
- $B^0 \to \mathrm{D D_s^-} X$; $\mathrm{D} \to \mu\nu X$: $5.7 \pm 1.9\%$.

The reconstruction efficiencies did not include any lifetime biasing cuts at this point. We, therefore, determined the efficiency of the lifetime selections for the different samples as a function of VPDL. Figure 4 shows the efficiency as a function of VPDL for the decay $B_s^0 \to D_s^- \mu^+ X$. In addition to this, K-factors for all the above decays were obtained using Monte Carlo simulations. Figure 5 shows the K-factor distributions for the $B_s^0 \to D_s^- \mu^+ X$ semileptonic decays. As expected, the K-factors for $\mathrm{D_s^{-*}}, \mathrm{D_{s0}^{*-}}$ and $\mathrm{D_{s1}^{'-}}$ have lower mean values because more decay products are missing. Note that since the K-factors were defined as the ratio of transverse momenta, they can exceed unity.

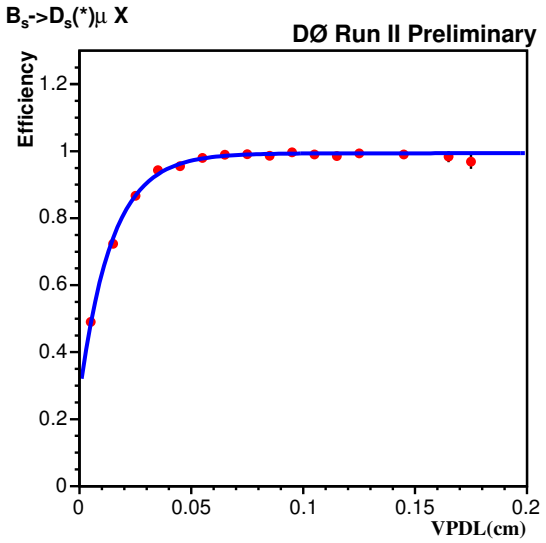

FIG. 4: Efficiency of the lifetime selection criteria as a function of VPDL (cm) for $B_s^0 \to D_s^- \mu^+ X$.

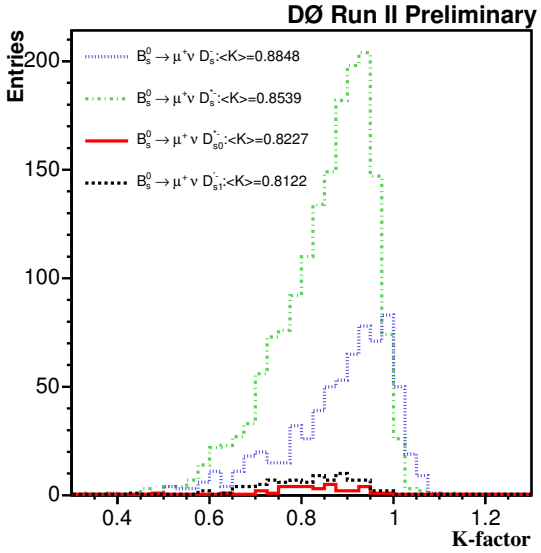

FIG. 5: K-factors for $\mathrm{B_s^0} \to \mu^+\nu \mathrm{D_s^-}$; $\mathrm{B_s^0} \to \mu^+\nu D_s^{*-} \to \mu^+\nu \mathrm{D_s^-}$; $\mathrm{B_s^0} \to \mu^+\nu D_{s0}^{*-} \to \mu^+\nu \mathrm{D_s^-}$; $\mathrm{B_s^0} \to \mu^+\nu D_{s1}^{'-} \to \mu^+\nu \mathrm{D_s^-}$ processes.

In addition to the above decays, $c\bar{c}$ pairs originating from gluon splitting can potentially contaminate the B_s^0 data sample. In principle tagging on the opposite side should suppress the contamination since a tag implies the presence of a b-quark on that side. The number of $c\bar{c}$ events in each VPDL bin were taken from the study done in Ref. [10] to be $3.5 \pm 2.5\%$ after tagging. A zero asymmetry was assigned to these events.

The decay length resolution for all the above samples was obtained using Monte Carlo simulations and was parameterized using Gaussians. Moreover, since simulations do not precisely model the uncertainties on the track parameters, a special procedure was developed to tune the track impact parameter resolution [10]. This procedure changed the track impact parameters and their errors in MC to produce a resolution similar to that in data. Signal MC was used to determine how this tuning procedure changed the VPDL resolution function.

Figure 6 shows the VPDL resolution for the decay $B_s^0 \to D_s^- \mu^+ X$ before and after tuning. The dashed blue line denotes the VPDL resolution from untuned Monte Carlo and is modeled using three Gaussians. The solid red line denotes the resolution from tuned Monte Carlo. The fractions and widths of the three Gaussians used to model the untuned MC were used as fixed parameters when fitting the VPDL resolution from tuned MC. The overall scale

factor was found to be equal to 1.168 ± 0.024. Additionally, since VPDL resolutions have been found to depend on the VPDL, a variable scale factor (that depends on VPDL resolution) was used to model this effect (Fig. 7). This dependence was incorporated into the asymmetry fitting procedure.

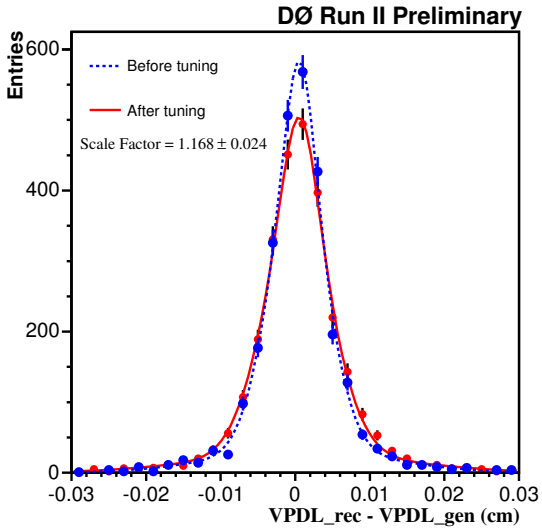

FIG. 6: VPDL resolution for the decay $B_s^0 \to D_s^- \mu^+ X$ before (dashed line) and after (solid line) tuning.

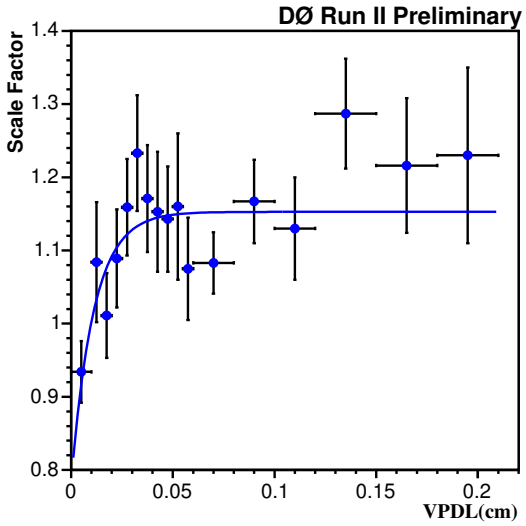

FIG. 7: The VPDL resolution scale factor as a function of VPDL. Bigger VPDL bins are used at larger values of VPDL owing to smaller statistics in those bins.

VII. MEASURED ASYMMETRY

The number of B_s^0 candidates in the flavor untagged sample is quite large and allows us to fit a large statistics sample. However, once the data is flavor tagged into mixed and unmixed samples and then separated into bins of VPDL the statistics in each bin are very much reduced. To improve on the fitting, we first fit the full untagged sample, and then fix the mass and width of the D_s^- from that sample when the flavor tagged samples are fit. Single Gaussians are used to describe the $D_s^- \to K^{*0}K^-$ and $D^- \to K^{*0}K^-$ decays, and the background is modeled by an exponential. The $D^- \to K^+\pi^-\pi^-$ reflection is modeled using two Gaussians as described earlier in the note, and the sigmas of the two Gaussians along with their relative fraction are used as fixed parameters. Figure 8 shows the measured asymmetry as a function of VPDL. The number of non-oscillated (unmixed) and oscillated (mixed) events along with the asymmetry for each VPDL bin are listed in Table I.

VIII. FITTING PROCEDURE FOR Δm_s LIMIT

Figure 8 shows that no B_s^0 oscillations can be resolved at the moment. Therefore, in order to set a lower limit on $\Delta \mathrm{m_s}$, Eq. 8 is modified to the following form:

$$n_s^{non-osc/osc}(x) = \frac{K}{c\tau_{B_s}} \exp\left(-\frac{Kx}{c\tau_{B_s}}\right) \cdot 0.5 \cdot [1 \pm (2\eta_s - 1)\cos(\Delta m_s \cdot Kx/c) \cdot \mathcal{A}] \tag{15}$$

where $\mathcal{A}$ is now the only fit parameter. Different choices of $\Delta \mathrm{m_s}$ in the range 1 ps^{-1} to 20 ps^{-1} are input and the fitted value of the amplitude, $\mathcal{A}$, is returned. By plotting the fitted value of $\mathcal{A}$ as a function of the input value of $\Delta \mathrm{m_s}$, one searches for a peak of $\mathcal{A}$=1 to obtain a measurement of $\Delta \mathrm{m_s}$. If no peak is found, limits can easily be set using this method. The sensitivity of a measurement is determined by calculating the probability that $\mathcal{A}$=0 could fluctuate to $\mathcal{A}$=1. This occurs as $1.645\sigma = 1$ (95% C.L.), where σ is the uncertainty associated with $\mathcal{A}$. The limit is determined by calculating the probability that a fitted value of $\mathcal{A}$ could fluctuate to $\mathcal{A} = 1$. This occurs at $\mathcal{A} + 1.645\sigma = 1$.

Figure 9 shows the dependence of the parameter $\mathcal{A}$ and its error on $\Delta \mathrm{m_s}$. A 95% confidence level limit on the oscillation frequency $\Delta \mathrm{m_s} > 5.1$ ps^{-1} and sensitivity 7.8 ps^{-1} were obtained with statistical error only.

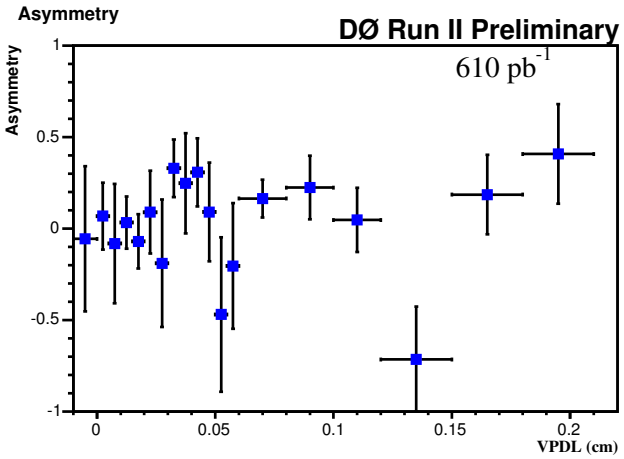

FIG. 8: The asymmetry in the $D_s^- \to K^{*0}K^-$ sample as a function of the visible proper decay length (VPDL).

TABLE I: For each of the 19 VPDL intervals the measured number of non-oscillated ($N_i^{non-osc}$) and oscillated (N_i^{osc}) D_s^- events, their statistical errors $\sigma(N_i^{non-osc})$ and $\sigma(N_i^{osc})$, the measured asymmetry, A_i, and its error $\sigma(A_i)$, all determined from the fits to the corresponding D_s^- mass distributions.

bin	VPDL range, cm	$N_i^{non-osc}$	$\sigma(N_i^{non-osc})$	N_i^{osc}	$\sigma(N_i^{osc})$	A_i	$\sigma(A_i)$
1	[−0.01, 0]	26.80	16.28	30.01	15.38	−0.056	0.396
2	[0, 0.005]	65.23	17.62	56.92	14.17	0.068	0.183
3	[0.005, 0.01]	34.75	18.95	40.99	15.08	−0.082	0.327
4	[0.01, 0.015]	83.21	18.67	77.99	13.71	0.032	0.142
5	[0.015, 0.02]	80.50	19.13	92.78	16.68	−0.071	0.148
6	[0.02, 0.025]	57.21	19.10	47.84	14.86	0.0892	0.226
7	[0.025, 0.03]	27.89	17.15	40.99	15.65	−0.190	0.349
8	[0.03, 0.035]	95.93	18.08	48.44	14.55	0.329	0.158
9	[0.035, 0.04]	49.17	16.68	29.69	14.08	0.247	0.274
10	[0.04, 0.045]	75.29	16.42	39.94	13.94	0.307	0.186
11	[0.045, 0.05]	39.81	14.79	33.24	13.19	0.090	0.270
12	[0.05, 0.055]	12.56	13.16	34.82	9.64	−0.470	0.422
13	[0.055, 0.06]	20.20	13.05	30.61	9.45	−0.205	0.343
14	[0.06, 0.08]	178.99	23.74	128.69	21.14	0.163	0.103
15	[0.08, 0.1]	90.90	20.84	57.68	16.46	0.224	0.174
16	[0.1, 0.12]	63.54	16.43	57.91	13.78	0.046	0.175
17	[0.12, 0.15]	12.70	14.80	76.76	16.08	−0.716	0.288
18	[0.15, 0.18]	43.96	12.07	30.24	10.77	0.185	0.217
19	[0.18, 0.21]	36.49	8.94	15.39	9.30	0.407	0.272

IX. SYSTEMATIC UNCERTAINTIES

All studied contributions to the systematic uncertainty of the amplitude are listed in Tables II and III. Some of the main contributions are due to uncertainties in resolution scale factor, $c\bar{c}$ contamination, $B_s^0 \to \mu D_s X$ branching ratio, K-factors, and the mass fitting procedure. The contribution of the different sources to the total systematic error was estimated using the following formula and summing in quadrature [6]:

$$\sigma_{\mathcal{A}}^{sys} = \Delta\mathcal{A} + (1 - \mathcal{A})\frac{\Delta\sigma_{\mathcal{A}}}{\sigma_{\mathcal{A}}}. \tag{16}$$

The result is shown in Fig. 9.

TABLE II: Systematic uncertainties on the amplitude for the range Δm_s = 1 ps^{-1}- 10 ps^{-1}. The shifts of both the measured amplitude, $\Delta\mathcal{A}$, and its statistical uncertainty, $\Delta\sigma$, are listed.

Osc. frequency		1 ps^{-1}	2 ps^{-1}	3 ps^{-1}	4 ps^{-1}	5 ps^{-1}	6 ps^{-1}	7 ps^{-1}	8 ps^{-1}	9 ps^{-1}	10 ps^{-1}
$\mathcal{A}$		0.162	-0.395	-0.292	-0.546	0.239	0.794	0.316	-0.343	-0.401	-0.331
Stat. uncertainty		0.228	0.269	0.299	0.419	0.428	0.455	0.508	0.637	0.773	0.900
$\eta_s = 0.684$	$\Delta\mathcal{A}$	+0.005	-0.016	-0.013	-0.021	+0.013	+0.035	+0.016	-0.010	-0.013	-0.012
	$\Delta\sigma$	+0.010	+0.012	+0.013	+0.019	+0.019	+0.020	+0.022	+0.028	+0.034	+0.039
$c\bar{c}:$ 6%	$\Delta\mathcal{A}$	-0.000	-0.046	-0.025	-0.066	+0.018	+0.094	+0.037	-0.058	-0.059	-0.038
	$\Delta\sigma$	+0.012	+0.011	+0.010	+0.022	+0.024	+0.029	+0.035	+0.052	+0.065	+0.070
Br(D_sD_s) = 4.7%	$\Delta\mathcal{A}$	-0.004	+0.011	+0.010	+0.015	-0.009	-0.023	-0.011	+0.006	+0.008	+0.008
	$\Delta\sigma$	-0.007	-0.009	-0.010	-0.013	-0.013	-0.014	-0.015	-0.019	-0.023	-0.027
Br($D_s\mu X$) = 5.5%	$\Delta\mathcal{A}$	+0.028	-0.032	-0.022	-0.057	-0.007	+0.041	-0.004	-0.062	-0.065	-0.047
	$\Delta\sigma$	+0.015	+0.018	+0.021	+0.028	+0.029	+0.030	+0.033	+0.041	+0.050	+0.059
$c\tau_{B_s} = 455\mu m$	$\Delta\mathcal{A}$	+0.002	+0.001	-0.002	-0.005	+0.005	+0.006	-0.004	-0.011	-0.006	-0.005
	$\Delta\sigma$	+0.000	+0.001	+0.000	+0.002	+0.002	+0.002	+0.003	+0.005	+0.006	+0.007
$\Delta\Gamma/\Gamma = 0.2$	$\Delta\mathcal{A}$	+0.001	-0.001	-0.002	-0.001	+0.001	-0.000	+0.000	+0.002	-0.000	-0.001
	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.000	+0.000	+0.000	+0.001
D_s mass changed to	$\Delta\mathcal{A}$	+0.002	+0.019	+0.008	+0.017	-0.016	-0.032	-0.026	-0.022	-0.028	-0.022
$M_s + 1\sigma$	$\Delta\sigma$	-0.004	-0.004	-0.005	-0.008	-0.008	-0.008	-0.010	-0.014	-0.017	-0.019
D_s width changed to	$\Delta\mathcal{A}$	-0.005	+0.003	+0.000	+0.004	-0.001	+0.007	+0.008	+0.007	+0.003	+0.002
$\sigma_s - 1\sigma$	$\Delta\sigma$	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001
$K^*\pi$ G1 width	$\Delta\mathcal{A}$	+0.012	-0.004	+0.002	-0.006	-0.018	-0.017	-0.013	-0.005	+0.022	+0.042
decreased 1σ	$\Delta\sigma$	-0.003	-0.004	-0.004	-0.006	-0.005	-0.005	-0.004	-0.005	-0.006	-0.008
$K^*\pi$ G2/G1 frac	$\Delta\mathcal{A}$	-0.024	-0.061	-0.019	-0.039	+0.045	+0.077	+0.060	+0.052	+0.017	-0.044
decreased 1σ	$\Delta\sigma$	+0.015	+0.018	+0.020	+0.033	+0.029	+0.028	+0.028	+0.038	+0.049	+0.057
D+ ratio	$\Delta\mathcal{A}$	-0.007	-0.039	-0.023	-0.014	+0.069	+0.073	+0.043	+0.036	+0.035	+0.026
increased by 25%	$\Delta\sigma$	+0.011	+0.013	+0.014	+0.024	+0.021	+0.020	+0.020	+0.025	+0.034	+0.039
mass bin size	$\Delta\mathcal{A}$	+0.008	-0.015	+0.037	+0.044	+0.050	+0.062	+0.029	+0.008	-0.023	-0.085
smaller by 25%	$\Delta\sigma$	-0.001	-0.000	-0.001	-0.000	-0.001	-0.000	-0.001	-0.005	-0.007	-0.011
Linear background	$\Delta\mathcal{A}$	-0.022	-0.002	+0.011	+0.035	-0.068	-0.044	-0.037	-0.076	-0.160	-0.171
	$\Delta\sigma$	-0.006	-0.005	-0.007	-0.010	-0.009	-0.010	-0.008	+0.001	-0.008	-0.011
Using corrected	$\Delta\mathcal{A}$	-0.015	-0.027	+0.040	+0.040	-0.039	+0.020	+0.075	+0.137	+0.154	+0.060
slopes for bkg	$\Delta\sigma$	-0.013	-0.018	-0.018	-0.025	-0.020	-0.017	-0.013	-0.019	-0.032	-0.033
same eff.Vs VPDL dependence	$\Delta\mathcal{A}$	+0.000	-0.005	-0.004	-0.005	-0.003	-0.001	-0.001	-0.001	-0.001	-0.001
for signal and bkg	$\Delta\sigma$	+0.001	+0.001	+0.002	+0.002	+0.002	+0.002	+0.002	+0.001	+0.002	+0.002
slope of eff.curve for signal	$\Delta\mathcal{A}$	+0.000	-0.002	-0.002	-0.004	-0.001	+0.001	-0.001	-0.005	-0.008	-0.009
changed 1 σ	$\Delta\sigma$	+0.000	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	+0.002	+0.003	+0.003
Resolution $S.F. = 1.192$	$\Delta\mathcal{A}$	+0.003	-0.004	-0.007	-0.042	+0.003	+0.071	+0.029	-0.080	-0.121	-0.125
	$\Delta\sigma$	+0.001	+0.003	+0.006	+0.017	+0.022	+0.031	+0.042	+0.069	+0.106	+0.139
Resolution $S.F. = 2$	$\Delta\mathcal{A}$	-0.008	-0.002	-0.003	+0.010	+0.008	-0.002	+0.004	+0.017	+0.017	+0.012
for background	$\Delta\sigma$	-0.001	-0.000	-0.000	-0.002	-0.003	-0.005	-0.006	-0.008	-0.010	-0.012
K-factor variation	$\Delta\mathcal{A}$	+0.016	-0.022	+0.021	-0.015	-0.106	+0.003	+0.102	+0.077	-0.028	-0.009
2%	$\Delta\sigma$	+0.000	-0.004	-0.001	-0.007	-0.006	-0.002	-0.010	-0.024	-0.017	-0.031
K-factor using	$\Delta\mathcal{A}$	-0.000	+0.002	+0.000	+0.001	+0.003	+0.001	-0.005	-0.006	+0.008	+0.004
reco values	$\Delta\sigma$	-0.000	-0.000	-0.000	-0.000	-0.001	-0.001	-0.001	-0.001	-0.002	-0.002
Using smoothed	$\Delta\mathcal{A}$	-0.000	-0.000	+0.000	+0.000	-0.000	+0.000	+0.001	+0.001	-0.001	-0.001
K-factor histograms	$\Delta\sigma$	-0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.000	+0.001
Total syst.	σ_{tot}^{sys}	0.142	0.163	0.133	0.161	0.249	0.218	0.223	0.234	0.285	0.276
Total (stat.+ syst.)	σ_{tot}	0.269	0.314	0.327	0.449	0.495	0.504	0.555	0.678	0.824	0.942

TABLE III: Systematic uncertainties on the amplitude for the range Δm_s = 11 ps^{-1} - 20 ps^{-1}. The shifts of both the measured amplitude, $\Delta\mathcal{A}$, and its statistical uncertainty, $\Delta\sigma$, are listed.

Osc. frequency		11 ps^{-1}	12 ps^{-1}	13 ps^{-1}	14 ps^{-1}	15 ps^{-1}	16 ps^{-1}	17 ps^{-1}	18 ps^{-1}	19 ps^{-1}	20 ps^{-1}
$\mathcal{A}$		-0.473	-0.083	0.750	0.886	-0.134	-0.645	-0.130	1.141	2.117	1.524
Stat. uncertainty		1.162	1.468	1.785	2.261	2.713	2.934	3.126	3.425	3.982	4.967
$\eta_s = 0.684$	$\Delta\mathcal{A}$	-0.020	-0.008	+0.022	+0.024	-0.016	-0.032	-0.003	+0.049	+0.092	+0.074
	$\Delta\sigma$	+0.051	+0.064	+0.078	+0.098	+0.118	+0.128	+0.116	+0.149	+0.191	+0.240
$c\bar{c}$: 6%	$\Delta\mathcal{A}$	-0.072	-0.022	+0.100	+0.128	-0.064	-0.181	-0.066	+0.237	+0.534	+0.594
	$\Delta\sigma$	+0.097	+0.124	+0.144	+0.201	+0.277	+0.319	+0.326	+0.389	+0.469	+0.687
Br(D_sD_s) = 4.7%	$\Delta\mathcal{A}$	+0.013	+0.005	-0.013	-0.015	+0.010	+0.019	+0.005	-0.033	-0.060	-0.047
	$\Delta\sigma$	-0.035	-0.044	-0.053	-0.067	-0.080	-0.087	-0.110	-0.101	-0.118	-0.137
Br($D_s\mu X$) = 5.5%	$\Delta\mathcal{A}$	-0.039	+0.027	+0.136	+0.174	+0.080	-0.006	+0.006	+0.083	+0.137	+0.038
	$\Delta\sigma$	+0.077	+0.097	+0.119	+0.151	+0.181	+0.194	+0.187	+0.227	+0.265	+0.346
$c\tau_{B_s} = 455\mu m$	$\Delta\mathcal{A}$	-0.005	+0.005	+0.018	+0.013	-0.008	-0.014	+0.003	+0.023	+0.036	+0.017
	$\Delta\sigma$	+0.010	+0.011	+0.013	+0.018	+0.021	+0.021	+0.005	+0.030	+0.039	+0.070
$\Delta\Gamma/\Gamma = 0.2$	$\Delta\mathcal{A}$	+0.001	+0.000	-0.001	-0.002	+0.001	+0.001	+0.005	+0.000	-0.002	-0.001
	$\Delta\sigma$	+0.001	+0.001	+0.001	+0.001	+0.001	+0.001	-0.018	+0.001	+0.001	+0.002
D_s mass changed to	$\Delta\mathcal{A}$	-0.012	-0.030	-0.070	-0.074	+0.006	+0.093	+0.126	+0.108	+0.088	+0.104
$M_s + 1\sigma$	$\Delta\sigma$	-0.027	-0.034	-0.041	-0.057	-0.075	-0.079	-0.097	-0.081	-0.088	-0.093
D_s width changed to	$\Delta\mathcal{A}$	+0.002	+0.002	+0.003	+0.005	+0.007	+0.010	+0.012	+0.011	+0.013	+0.019
$\sigma_s - 1\sigma$	$\Delta\sigma$	-0.001	-0.000	+0.000	+0.000	-0.001	-0.002	-0.022	-0.005	-0.007	-0.002
$K^*\pi$ G1 width	$\Delta\mathcal{A}$	+0.044	+0.010	-0.028	-0.063	-0.085	-0.107	-0.130	-0.174	-0.227	-0.294
decreased 1σ	$\Delta\sigma$	-0.013	-0.019	-0.026	-0.037	-0.041	-0.034	-0.048	-0.030	-0.033	-0.033
$K^*\pi$ G2/G1 frac	$\Delta\mathcal{A}$	-0.100	-0.035	+0.151	+0.290	+0.208	+0.074	+0.057	+0.139	+0.267	+0.378
decreased 1σ	$\Delta\sigma$	+0.082	+0.114	+0.148	+0.212	+0.261	+0.247	+0.208	+0.230	+0.263	+0.272
D+ ratio	$\Delta\mathcal{A}$	-0.013	-0.058	-0.046	-0.000	+0.011	+0.030	+0.061	+0.107	+0.162	+0.182
increased by 25%	$\Delta\sigma$	+0.058	-0.054	+0.106	+0.151	+0.176	+0.159	+0.126	+0.150	+0.172	+0.224
mass bin size	$\Delta\mathcal{A}$	-0.201	-0.335	-0.401	-0.402	-0.217	-0.081	-0.127	-0.250	-0.330	-0.353
smaller by 25%	$\Delta\sigma$	-0.018	-0.031	-0.044	-0.073	-0.101	-0.094	-0.095	-0.065	-0.061	-0.055
Linear background	$\Delta\mathcal{A}$	-0.143	-0.113	-0.059	+0.107	+0.357	+0.465	+0.489	+0.538	+0.601	+0.685
	$\Delta\sigma$	-0.019	-0.031	-0.041	-0.048	-0.037	-0.024	-0.040	-0.026	-0.029	-0.015
Using corrected	$\Delta\mathcal{A}$	-0.030	-0.054	-0.008	-0.182	-0.438	-0.479	-0.409	-0.382	-0.432	-0.571
slopes for bkg	$\Delta\sigma$	-0.038	-0.034	-0.044	-0.080	-0.107	-0.101	-0.106	-0.083	-0.096	-0.203
same eff.Vs VPDL dependence	$\Delta\mathcal{A}$	+0.001	+0.003	+0.005	+0.005	+0.008	+0.010	+0.010	+0.001	-0.005	-0.001
for signal and bkg	$\Delta\sigma$	+0.002	+0.003	+0.004	+0.003	+0.001	-0.001	-0.020	-0.002	-0.004	+0.002
slope of eff.curve for signal	$\Delta\mathcal{A}$	-0.014	-0.016	-0.018	-0.032	-0.038	-0.033	-0.042	-0.081	-0.128	-0.158
changed 1 σ	$\Delta\sigma$	+0.005	+0.005	+0.003	-0.003	-0.016	-0.015	-0.051	-0.040	-0.056	-0.084
Resolution $S.F.$ = 1.192	$\Delta\mathcal{A}$	-0.236	-0.210	-0.004	-0.070	-0.607	-0.919	-0.786	-0.273	+0.063	-0.492
	$\Delta\sigma$	+0.215	+0.310	+0.391	+0.529	+0.687	+0.807	+0.907	+1.075	+1.246	+1.474
Resolution $S.F.$ = 2	$\Delta\mathcal{A}$	+0.013	-0.002	-0.030	-0.040	-0.015	+0.008	+0.013	-0.005	-0.009	+0.029
for background	$\Delta\sigma$	-0.016	-0.021	-0.026	-0.033	-0.039	-0.043	-0.063	-0.048	-0.049	-0.038
K-factor variation	$\Delta\mathcal{A}$	+0.026	-0.159	-0.169	+0.082	+0.302	+0.037	-0.265	-0.430	-0.220	+0.424
2%	$\Delta\sigma$	-0.064	-0.062	-0.089	-0.135	-0.105	-0.041	-0.080	-0.123	-0.239	-0.423
K-factor using	$\Delta\mathcal{A}$	-0.009	+0.004	+0.016	-0.004	-0.016	-0.035	+0.015	+0.065	+0.048	-0.008
reco values	$\Delta\sigma$	-0.002	-0.004	-0.005	-0.005	-0.007	-0.010	-0.028	-0.008	-0.006	+0.025
Using smoothed	$\Delta\mathcal{A}$	+0.003	-0.006	-0.003	-0.012	+0.027	-0.003	+0.019	-0.009	-0.037	+0.015
K-factor histograms	$\Delta\sigma$	+0.001	+0.001	+0.002	+0.002	+0.002	+0.002	-0.016	+0.002	+0.003	+0.012
Total syst.	σ_{tot}^{sys}	0.323	0.477	0.533	0.612	0.859	0.906	0.906	0.956	1.017	1.441
Total (stat.+ syst.)	σ_{tot}	1.206	1.543	1.863	2.342	2.845	3.070	3.254	3.556	4.110	5.172

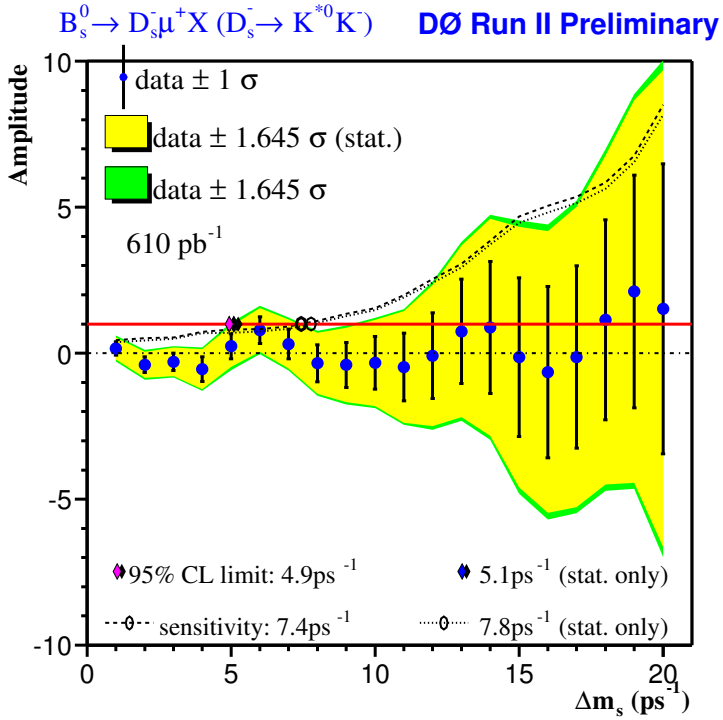

FIG. 9: B_s^0 oscillation amplitude with statistical and systematic errors ($B_s^0 \to D_s^- \mu^+ X$; $D_s^- \to K^{*0} K^-$).

X. COMBINATION WITH $B_s^0 \to D_s^- \mu^+ X$ ($D_s^- \to \phi\pi^-$)

Figure 10 shows the dependence of the parameter $\mathcal{A}$ and its error on Δm_s for the $D_s^- \to K^{*0} K^-$ and $D_s^- \to \phi\pi^-$ [4] combination. A common procedure [11] was used for combining the two results and the following contributions to the systematic error were considered as correlated for these analyses:

- Uncertainty in purity ["$\eta_s = 0.684$" in Tables II and III];
- Uncertainty in $c\bar{c}$ contamination ["$c\bar{c}:$ 6%" in Tables II and III];
- Uncertainty in $D_s D_s$ contribution ["Br$(D_s D_s) =$ 4.7%" in Tables II and III];
- Uncertainty in branching ratio $B_s^0 \to \mu D_s X$ ["Br$(D_s \mu X) =$ 5.5%" in Tables II and III];
- Uncertainty in B_s^0 lifetime ["$c\tau_{B_s^0} = 455 \mu m$" in Tables II and III];
- Non-zero $\Delta\Gamma_s/\Gamma_s$ ["$\Delta\Gamma/\Gamma = 0.2$" in Tables II and III].

A 95% confidence level limit on the oscillation frequency $\Delta m_s > 7.3$ ps^{-1} and sensitivity 9.5 ps^{-1} were obtained.

XI. CONCLUSIONS

The $B_s^0 \to D_s^- \mu^+ X$ decay into the $D_s^- \to K^{*0} K^-$ ($K^{*0} \to K^+ \pi^-$) final state was reconstructed at DØ using $\sim$ 610 pb^{-1} of data. A search for B_s^0 oscillations was performed using opposite side flavor tagging algorithms and a 95% confidence level limit on the oscillation frequency $\Delta m_s > 4.9$ ps^{-1} and a sensitivity of 7.4 ps^{-1} were obtained. A combination with the $B_s^0 \to D_s^- \mu^+ X$ ($D_s^- \to \phi\pi^-$) decay mode improves the result for the limit to $\Delta m_s > 7.3$ ps^{-1} and sensitivity to 9.5 ps^{-1}.

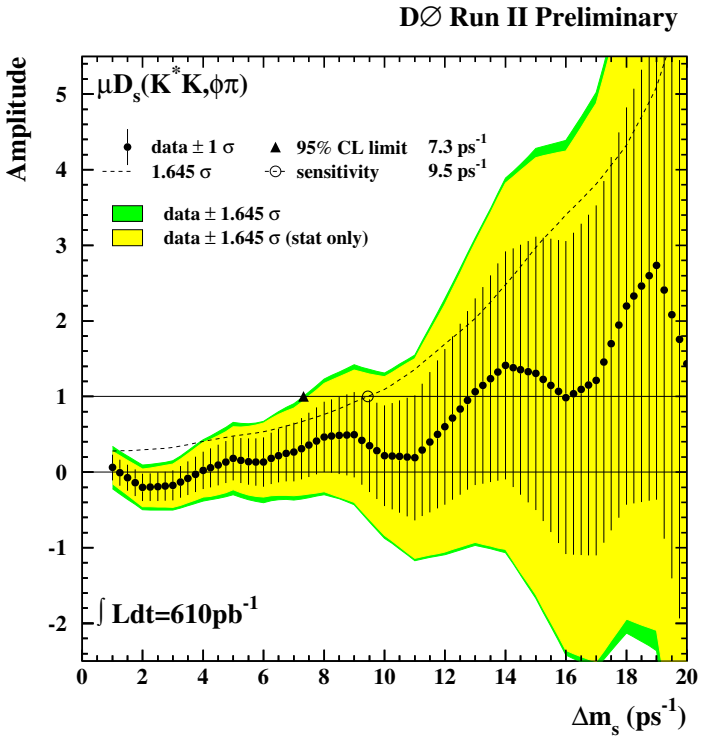

FIG. 10: Combined B_s^0 oscillation amplitude.

Acknowledgments

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CAPES, CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); PPARC (United Kingdom); MSMT (Czech Republic); CRC Program, CFI, NSERC and WestGrid Project (Canada); BMBF and DFG (Germany); SFI (Ireland); Research Corporation, Alexander von Humboldt Foundation, and the Marie Curie Program.

[1] H.Albrecht *et al.* (ARGUS Collaboration),"Observation of B0- Anti-B0 Mixing," Phys. Lett. **B192**, 245 (1987).
[2] The Review of Particle Physics by the Particle Data Group, S.Eidelman *et al.*, Phys. Lett. **B592**, 1 (2004).
[3] J. Charles *et al.* [CKMfitter Group], "CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories," arXiv:hep-ph/0406184.
[4] DØ Collaboration, "Mixing in the $B_s^0 - \bar{B}_s^0$ System Using $B_s^0 \to D_s^- \mu^+ X$,$D_s^- \to \phi\pi^-$ Decay Mode And Opposite-Side Flavor Tagging," DØ Note 4881.
[5] DØ Collaboration, "Combined Opposite-side Flavor Tagging," DØ Note 4875.
[6] H.G.Moser and A.Roussarie, Nucl. Instrum. Methods A **384**, 491 (1997).
[7] `http://www-d0.fnal.gov/computing/algorithms/muon/muon_algo.html`.
[8] S. Catani, Yu.L. Dokshitzer, M. Olsson, G. Turnock, B.R. Webber, Phys. Lett. **B269** (1991) 432.
[9] D.J. Lange, NIM A 462 (2001) 152; for details see `http://www.slac.stanford.edu/~ lange/EvtGen`.
[10] DØ Collaboration, "Mixing in the $B_s^0 - \bar{B}_s^0$ system using semileptonic decay modes and opposite side flavor tagging," DØ Note 4724.
[11] H. F. A. Group(HFAG), "Averages of b-hadron properties as of winter 2005," arXiv:hep-ex/0505100.
[12] Conjugate modes are implied throughout the note.
[13] In the plane perpendicular to the beam direction.
[14] Parallel to the beam direction.